INNOVATION IN LANGUAGE LEARNING International Conference

Language Teachers in the Age of Generative AI: A Contextualised Review of International Evidence and Implications for Teacher Education in Malta

Joseph Vancell

Faculty of Education, University of Malta, Malta

Abstract

The rapid diffusion of generative artificial intelligence (GenAI), particularly large language models, has distinctive implications for language education. This contextualised narrative review examines international and local research on language teachers and student teachers. The literature suggests that rather than just focusing on writing prompts, teachers and students get far more out of design guidance, real classroom practice, and critical reflection. However, most of the reviewed research focuses on English teaching outside of Malta, leans too heavily on self-reported survey data, and offers little proof of how these practices hold up in real classrooms over time. This review situates these limitations within the University of Malta's Master in Teaching and Learning (MTL), with particular attention to how the programme integrates educational inquiry, subject pedagogy, and field placement. The paper proposes a research agenda examining the practices and decision-making across planning, teaching, feedback, and supervision identified in the literature.

Keywords: *generative AI; language teacher education; pre-service teachers; professional judgement; Malta; field placement*

1. Introduction

By 2026, generative artificial intelligence (GenAI) tools had demonstrated considerable versatility across a broad spectrum of tasks. Their capabilities extended beyond generating text, images and music to include idea generation, data analysis and evaluation, summarisation, paraphrasing, translation, conversational interaction, sentiment analysis, text classification, named entity recognition, code generation and image captioning. A language teacher while preparing a lesson can now ask a large language model (LLM) to suggest a reading passage, simplify its vocabulary, devise comprehension questions and produce feedback on a pupil's writing. Much of this can happen within seconds. The issue now is not if the technology can generate teaching and learning content, but rather if the generated content is linguistically correct, provides educational value, and is appropriate for the students that will be using it. Student teachers face a similar issue in terms of what type of professional they are becoming. The earliest scholarship, understandably, focused on the capabilities of GenAI. Practical accounts illustrated how it could support classroom activities, materials development and assessment [2, 7, 5, 13]. Subsequent empirical work turned to teacher competence, professional judgement and the conditions under which apparently useful outputs translate into worthwhile teaching [16, 17, 20], while also engaging with cognitive and ethical concerns. This shift is significant for initial teacher education, since skill in crafting a prompt reveals little, in itself, about a teacher's understanding of language development, differentiation, assessment or classroom relationships. This paper offers a contextualised narrative review of that developing literature and considers its relevance to Malta. It asks: What do international and Maltese studies suggest about how language teachers learn to work with GenAI? Which claims are supported by classroom evidence, and which remain provisional? And what questions should be pursued within the University of Malta's Master in Teaching and Learning (MTL)? The argument is that teacher education should place guided design, classroom practice and critical reflection at the centre of GenAI-related professional learning. Prompting is part of the work, but it should not be mistaken for the work itself.

2. Review Methodology and Boundaries

The purpose of this review is to build an interpretive argument through the reading of published international and Malta-based research in the field of GenAI in language teaching and learning, including systematic and scoping reviews, intervention studies, case studies of pre-service and practising

teachers, and studies of language learning and professional practice. The review, considered papers indentified through Scopus, ERIC, JSTOR, Web of Science, Google Scholar and Hydi (the University of Malta's article database) from 2023 to 2026, with an emphasis on language teachers, student teachers, lesson planning, classroom practice, feedback, digital literacy and professional learning.
Sources were selected and analysed for their research context, design, participants, central claims and relevance to language teacher education. Particular attention was paid to the difference between reported perceptions, analyses of teaching artefacts, observed practice and measured learner outcomes. Findings were then used for a critical analysis of the MTL. The study makes no claim of an exhaustive database search, formal risk-of-bias assessment or independently reproduced effect-size calculation. As a result, the review cannot establish the prevalence of a practice across all studies or offer a pooled estimate of teacher-development outcomes. Its value lies instead in a transparent, context-sensitive synthesis that identifies credible directions for inquiry.

3. What the Growing Evidence Base Tells Us

Early reviews, most from 2023, mapped a growing research field. Law's scoping review included 41 publications and found substantial attention to English as a Foreign Language, writing and higher education, alongside calls for more empirical research and consideration of privacy and teacher AI literacy [9]. Li and colleagues reviewed 36 papers from the first year following ChatGPT's release, likewise identifying gaps in longitudinal work, feedback quality and diverse educational settings [11]. Later syntheses broadened the picture: Lee and colleagues concentrated on classroom-based language education [10], Kozlova synthesised 32 studies of foreign language education and pointed to uneven teacher preparedness [8], while Rejeb and colleagues used bibliometric and topic-modelling methods to map 908 publications from 2023–2025 [18]. The growing volume of publications signals an active field; it does not, however, settle whether classroom learning has improved.
There is promising quantitative evidence. In a task-based language teaching meta-analysis, Li, Shadiev and Chiu synthesised 25 studies involving 2,431 learners and reported a positive overall effect on linguistic performance (Hedges's $g = 0.75$). However, heterogeneity was exceptionally high ($I^2 = 93.89\%$), and outcomes varied with factors including learners' AI literacy, task planning and who assessed the work [12]. The finding therefore supports qualified optimism rather than the conclusion that any GenAI activity will benefit any learner. The educational design of the task, and the continuing involvement of a teacher, remain consequential.
A further distinction cuts across these reviews: many publications discuss potential applications, perceptions or short-term responses; fewer follow teachers through sustained lesson preparation, classroom interaction and revision. The question for teacher education is not simply whether a tool performs well in isolation, but how teachers learn to notice where it fails, justify its use and take responsibility for what follows.

4. Exercising Professional Judgement

Evidence on teachers' preparedness is revealing. In interviews with 12 university language instructors in Hong Kong, Kohnke, Moorhouse and Zou found that participants were curious about GenAI yet unsure how to use it in the classroom, what it meant for academic integrity and what professional support they would need to teach with it responsibly [6]. In a multinational qualitative study, Al-Khresheh recorded similar tensions among 46 English teachers: anticipated personalisation was accompanied by concerns about linguistic reliability, dependence and the neglect of speaking and listening [1]. Such studies matter because they reveal how teachers explain their concerns. Yet, as accounts of perceptions, they cannot establish what happens in lessons after the interview ends.
Evidence from initial teacher education takes the discussion further. Moorhouse and colleagues examined an 11-week intervention with 54 Master of Education students in Hong Kong, designed to develop their professional competence in GenAI [16]. The authors reported encouraging development, particularly in pedagogical competence and critical awareness, while noting that growth was uneven across the competence framework. This suggests that purposeful education can make a difference, without demonstrating that a single intervention produces durable classroom practice after graduation.
Two studies expose the limits of equating competence with prompting. Kerr and Kim found that South Korean pre-service EFL teachers used GenAI in selecting topics, producing materials, structuring lessons and checking language, but struggled with output quality and the integration of sound teaching strategies [4]. Moorhouse, Ho, Wu and Wan examined artefacts, interaction logs and reflections from 25 pre-service teachers and described a process of task identification, iterative prompting and task

implementation [15]. The significant move was from an attractive answer on a screen towards a defensible pedagogical decision. Teachers needed subject-pedagogical knowledge to know what to request, what to reject and why.

A 2026 study by Moorhouse, Hafner and Ho adds another dimension. Ten groups of pre-service TESOL teachers developed multimodal websites and used GenAI selectively, weighing convenience against authenticity, audience, creativity and the wish to remain meaningfully involved in the work [14]. Deliberate restraint could be as thoughtful as tool use. This perspective accords with Moorhouse and Wong's emphasis on evidence-informed, ethical and critical professional competence [17]. It also reminds us that language education concerns meaningful communication, not only efficient text production.

The strongest bridge to practice comes from research that follows teachers beyond initial enthusiasm. Wan and Moorhouse observed and interviewed ten university English instructors participating in collaborative lesson study in China. GenAI became entwined with planning, instruction and assessment, yet the continuation of new practices depended on perceived utility, professional risk and situated professional relationships [20]. Here digital literacy emerges less as a checklist of technical skills than as a social practice shaped by context. A teacher may be capable of using a tool and still have good reasons not to use it in a particular classroom.

5. Classroom Use, Feedback and Ethical Responsibility

Practical contributions by Bonner and colleagues, Koraishi, and Kohnke and colleagues show why language teachers might wish to experiment: GenAI can offer alternative texts, simulated exchanges, draft feedback and differentiated exercises [2, 7, 5]. Still, a workable activity on paper is not the same as a lesson carried out in the classroom. Teaching requires the teacher to listen to students, notice confusion, adapt the task and judge whether assistance is supporting language development or quietly replacing it. The distinction is especially sharp in writing, where a polished final product can conceal how little composing, revising or decision-making the learner actually undertook.

Assessment is therefore a test case for professional judgement. Teachers must decide which aspects of an assignment should remain demonstrably the learner's own, where AI-generated suggestions may function as formative scaffolding, how learners should acknowledge assistance and what evidence of process should be retained. A useful formative arrangement might ask learners to critique competing AI suggestions, explain why one revision improves meaning, and retain an annotated sequence of drafts. This is offered as a proposed instructional response, not as evidence that the literature has already validated this particular arrangement in Maltese classrooms.

Ethical questions are inseparable from pedagogy. Law identifies uneven attention to data privacy, security and AI literacy [9]; Kozlova records concerns about academic integrity and institutional support [8]. These issues become concrete when a student teacher considers uploading identifiable pupil work, asks an AI tool to assess a child, or uses generated cultural examples without checking their accuracy or appropriateness. Teacher education must therefore include decisions about what data not to share, what output not to trust, and when human feedback matters more than automated convenience.

6. The Maltese Context

Although international evidence is valuable, it cannot be assumed to transfer to the Maltese context. Malta brings together Maltese and English, additional foreign languages, varied learner repertoires and the particular vulnerability of a language less extensively represented in many AI training datasets. In a two-week study involving 41 international adult learners of Maltese, Żammit found that participants valued ChatGPT's accessibility but encountered limitations in grammar, vocabulary, conversational support and accuracy [24]. A tool appearing reliable in English cannot be presumed equally reliable in Maltese.

Żammit's later qualitative work with 47 secondary-school teachers of Maltese, including 40 reporting some GenAI use, describes experimentation with storytelling, contextual exercises and personalised feedback. Teachers perceived benefits for engagement and proficiency but these perceptions were not independently established as gains in learning outcomes [25]. A 2026 article explores related accounts of innovative practice with a sample of 47 teachers [26]. Because the sample descriptions are so similar, the two publications should not automatically be treated as independent replications without confirming participant overlap. Both nevertheless illuminate locally meaningful possibilities and the continuing importance of teacher preparation.

Research on English writing instruction in Malta contributes a complementary account of professional learning. Xerri and Block describe how a research–practice partnership enabled two teachers to reconsider writing instruction through joint inquiry, experimentation and reflection, rather than simply adopt AI-generated texts [21]. Their subsequent discussion of GenAI-oriented professional learning and a related chapter on 'liquid identities' extend that account towards collaboration, teacher agency and changing professional identities [22, 23]. These publications offer valuable depth, but, once again, should not be counted as multiple unrelated interventions without careful attention to their shared project origins.

Colman's study of English and Maltese speeches generated for conference-interpreter training offers a pointed local reminder: usable AI texts still required extensive post-editing, with weaknesses in argumentation and factual substance and more linguistic errors in Maltese [3]. Although interpreter training differs from school language teaching, the pedagogical lesson travels well: surface fluency should never replace scrutiny of communicative purpose, disciplinary quality or linguistic form. Taken together, the Malta-based studies support neither rejection nor uncritical adoption. They support careful, language-sensitive experimentation.

7. Implications for the University of Malta's MTL

The MTL provides a useful setting for examining these questions because it explicitly connects pedagogical content knowledge, educational context knowledge and research and inquiry with field placement in schools. Its stated purpose is to bridge theory and practice and prepare reflective, inquiring educators [19]. GenAI-related learning could be located within those existing connections rather than treated as a detached technical workshop. The suggestion here is conceptual: it does not imply that these activities are already established requirements of the programme.

One practical sequence could begin with a student teacher identifying a genuine language-learning problem from curriculum expectations and classroom observation. The student would then design a task, test AI assistance against disciplinary criteria, and explain what was retained, edited or discarded. During field placement, the task could be enacted where school policy, mentor agreement, consent and data safeguards allow. Afterwards, student and mentor could compare the intended lesson with what actually happened: who participated, what pupils understood, whether feedback helped, and what the next lesson should change. The process would create evidence of pedagogical reasoning, not merely evidence of tool use.

This sequence provides a natural point of convergence for language subject pedagogy, educational inquiry and field-placement supervision. It would also create room for a professionally significant decision: opting against GenAI where the learning objective, the learners' circumstances or the ethical risks do not justify its use. Put differently, teacher competence ought to be evident in the quality of judgement exercised, including the judgement to decline a tool.

8. A Research Agenda

Four connected questions warrant investigation within the Maltese MTL context. First, how do pre-service language teachers decide when GenAI is pedagogically appropriate during lesson planning, and how do those decisions differ across Maltese, English and other languages, including Italian? Second, how are AI-assisted plans changed, supported or abandoned during actual classroom teaching? Third, how do student teachers use, question and disclose GenAI when producing materials, assessing learners and giving feedback? Fourth, how do field-placement mentors and university supervisors recognise, challenge and develop student teachers' professional judgement about these practices?

A longitudinal, practice-focused design would yield richer insight than a single end-of-course questionnaire on perceptions. Consenting student teachers could be followed across one or both placement cycles using lesson plans and successive drafts, optional and appropriately anonymised AI interaction records, reflective accounts, classroom observation, mentor feedback and post-lesson interviews. Where feasible, learner work could be examined in de-identified form to investigate what feedback was received and how it was used. Sampling should deliberately include different languages and school contexts rather than treating an English-medium lesson as representative of all language teaching in Malta.

Analysis could trace decision points: why a prompt was written, why an output was trusted or rejected, how a task changed when pupils responded unexpectedly, and whether the student teacher's reasoning developed across both placements. Cross-case comparison would identify common practices alongside differences associated with language, school setting, mentor support and institutional policy. Ethical

safeguards would be central: participation must be voluntary and separate from assessment; supervisors' dual roles require careful management; pupils and schools must not be identifiable; and commercially hosted tools must not receive personal learner data without appropriate authorisation. These are design recommendations, not descriptions of an approved or completed empirical study.
The contribution of such research would be modest but consequential. Rather than ask whether future teachers like GenAI, it would examine how they use or refuse it, what happens when their plans meet a real classroom, and whether they can explain their choices in language-pedagogical terms. That is the level at which confidence becomes competence and experimentation has a chance of becoming responsible practice.

9. Conclusion

GenAI is expanding rapidly and so is research on its pedagogic use. The technology has made it easier to produce language-teaching resources. However, it has not made the professional task of deciding what learners need any simpler. International research increasingly points towards guided, situated and reflective teacher learning, while reviews remind us that much evidence remains English-centred, short term or dependent on reports of perceived benefit. Malta-based studies sharpen the argument by foregrounding the linguistic limits of AI-generated Maltese, the value of collaborative inquiry and the need to inspect apparently fluent materials critically.
For the MTL, the next question should not be how quickly student teachers can learn to write better prompts for language teaching and learning in general. It should be how they develop the judgement to plan with GenAI, teach with or without it, assess its consequences and explain their decisions to learners, mentors and supervisors. A research agenda grounded in the field placement could bring these often-hidden decisions into view and offer a more lasting foundation for teacher education than technical proficiency alone.

10. Declaration of Generative Ai and AI-assisted Technologies in the Writing Process

During the preparation of this work, the author used Google Gemini and NotebookLM in order to improve readability and language. After using these tools/services, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

REFERENCES

[1] Al-khresheh, M. H. (2024). Bridging technology and pedagogy from a global lens: Teachers' perspectives on integrating ChatGPT in English language teaching. *Computers and Education: Artificial Intelligence, 6*, 100218. https://doi.org/10.1016/j.caeai.2024.100218

[2] Bonner, E., Lege, R., & Frazier, E. (2023). Large language model-based artificial intelligence in the language classroom: Practical ideas for teaching. *Teaching English with Technology, 23*(1), 23–41. https://doi.org/10.56297/BKAM1691/WIEO1749

[3] Colman, A. (2026). GenAI for speech writing in the training of Maltese conference interpreters for the European Union. *International Journal of Language, Translation and Intercultural Communication, 11*, 43–70. https://doi.org/10.12681/ijltic.44157

[4] Kerr, R. C., & Kim, H. (2025). From prompts to plans: A case study of pre-service EFL teachers' use of generative AI for lesson planning. *English Teaching, 80*(1), 95–118. https://doi.org/10.15858/engtea.80.1.202503.95

[5] Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal, 54*(2), 537–550. https://doi.org/10.1177/00336882231162868

[6] Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). Exploring generative artificial intelligence preparedness among university language instructors: A case study. *Computers and Education: Artificial Intelligence, 5*, 100156. https://doi.org/10.1016/j.caeai.2023.100156

[7] Koraishi, O. (2023). Teaching English in the age of AI: Embracing ChatGPT to optimize EFL materials and assessment. *Language Education and Technology, 3*(1), 55–72. https://langedutech.com/letjournal/index.php/let/article/view/48

[8] Kozlova, Y. (2026). Generative AI in foreign language education: A scoping review of pedagogical applications and challenges. *Applied Language Learning, 35*(1), 13–40. https://www.dliflc.edu/ojs/ALL/article/view/67

[9] Law, L. (2024). Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review. *Computers and Education Open, 6*, 100174. https://doi.org/10.1016/j.caeo.2024.100174
[10] Lee, S., Choe, H., Zou, D., & Jeon, J. (2026). Generative AI (GenAI) in the language classroom: A systematic review. *Interactive Learning Environments, 34*(1), 335–359. https://doi.org/10.1080/10494820.2025.2498537
[11] Li, B., Lowell, V. L., Wang, C., & Li, X. (2024). A systematic review of the first year of publications on ChatGPT and language education: Examining research on ChatGPT's use in language learning and teaching. *Computers and Education: Artificial Intelligence, 7*, 100266. https://doi.org/10.1016/j.caeai.2024.100266
[12] Li, Y., Shadiev, R., & Chiu, T. K. F. (2026). Applying generative artificial intelligence to task-based language teaching and learning: A systematic review and meta-analysis. *TechTrends, 70*, 150–163. https://doi.org/10.1007/s11528-025-01140-7
[13] Moorhouse, B. L. (2024). Generative artificial intelligence and ELT. *ELT Journal, 78*(4), 378–392. https://doi.org/10.1093/elt/ccae032
[14] Moorhouse, B. L., Hafner, C. A., & Ho, T. Y. (2026). GenAI-assisted digital multimodal composing in TESOL teacher education. *TESOL Quarterly*. Advance online publication. https://doi.org/10.1002/tesq.70184
[15] Moorhouse, B. L., Ho, T. Y., Wu, C., & Wan, Y. (2025). Pre-service language teachers' task-specific large language model prompting practices. *RELC Journal*. Advance online publication. https://doi.org/10.1177/00336882251313701
[16] Moorhouse, B. L., Wan, Y., Wu, C., Kohnke, L., Ho, T. Y., & Kwong, T. (2024). Developing language teachers' professional generative AI competence: An intervention study in an initial language teacher education course. *System, 125*, 103399. https://doi.org/10.1016/j.system.2024.103399
[17] Moorhouse, B. L., & Wong, K. M. (2025). *Generative artificial intelligence and language teaching*. Cambridge University Press. https://doi.org/10.1017/9781009618823
[18] Rejeb, A., Rejeb, K., Zaher, H. F., Hassoun, A., Lahbib, M., & Süle, E. (2026). The impact of generative artificial intelligence on language teaching and learning. *Quality & Quantity*. Advance online publication. https://doi.org/10.1007/s11135-026-03068-3
[19] University of Malta. (2026). *Master in Teaching and Learning (MTL) in English (one main area): Course overview, 2026–2027*. https://www.um.edu.mt/courses/overview/pmtlengftt3-2026-7-o/
[20] Wan, Y., & Moorhouse, B. L. (2026). Generative AI as a catalyst: Transforming digital literacy practices in English language teaching. *Language Teaching Research*. Advance online publication. https://doi.org/10.1177/13621688261487631
[21] Xerri, D., & Block, J. (2024). Learning to be a writing instructor in the age of generative AI: The role of research-practice partnerships. *Language Teacher Education Research, 1*, 64–78. https://doi.org/10.32038/lter.2024.01.04
[22] Xerri, D., & Block, J. (2025). Unlocking teachers' generative AI-oriented professional learning through research–practice partnerships. *Teacher Development Academic Journal, 5*(1), 69–81. https://www.danielxerri.com/uploads/4/5/3/0/4530212/tdaj_v5i1_xerri___block.pdf
[23] Xerri, D., Block, J., Boezio, J. L., & Davis, S. (2026). Liquid identities and generative AI: A research-practice partnership exploring English language teachers' integration of AI tools for teaching writing. In S. W. Chong, H. Reinders, & Q. Liu (Eds.), *Research-practice partnership in language education* (pp. 185–206). Palgrave Macmillan. https://doi.org/10.1007/978-3-032-17518-2_8
[24] Żammit, J. (2024). Capturing the full potential of Maltese language learning through ChatGPT. *Technology in Language Teaching & Learning, 6*(1), 1–22. https://doi.org/10.29140/tltl.v6n1.1082
[25] Żammit, J. (2025). Secondary school teachers' experiences with generative AI in Maltese language teaching. *Technology, Knowledge and Learning, 30*, 2103–2121. https://doi.org/10.1007/s10758-025-09861-7
[26] Żammit, J. (2026). Applying the artificial intelligence muse: Secondary school teachers' innovative approaches in Maltese language teaching and learning. *Technology, Pedagogy and Education*. Advance online publication. https://doi.org/10.1080/1475939X.2026.2715526